

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-263279

(43)Date of publication of application : 19.09.2003

(51)Int.Cl.

G06F 3/06

G06F 12/08

(21)Application number : 2003-086908

(71)Applicant : HITACHI LTD

(22)Date of filing : 18.09.1998

(72)Inventor : FUJIMOTO KAZUHISA

TANAKA ATSUSHI

FUJIBAYASHI AKIRA

KANAI HIROKI

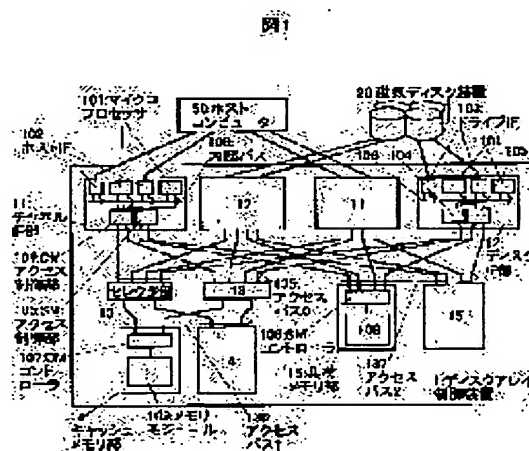
MINOWA NOBUYUKI

## (54) DISK ARRAY CONTROL APPARATUS

## (57)Abstract:

PROBLEM TO BE SOLVED: To provide a disk array control apparatus which realizes a high throughput and short response time with consideration of properties of data to be stored into cache memory and shared memory, and access properties for these memories.

SOLUTION: The disk array control apparatus has a plurality of channels IF, a plurality of disks IF, a cache memory and a shared memory. The connection form between the plurality of channels IF and the plurality of disks IF, and the cache memory is different from that between the plurality of channels IF and the shared memory, and the plurality of disks IF and the shared memory. In this disk array control apparatus, access paths to the cache memory and the shared memory are increased to improve throughput, and on the other hand, access time to the shared memory can be shortened.



## LEGAL STATUS

[Date of request for examination]

25.08.2005

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]



## 【特許請求の範囲】

【請求項 1】 ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報を格納する共有メモリ部とを有し、各チャネルインターフェース部は、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、各ディスクインターフェース部は、前記磁気ディスク装置とのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行するディスクアレイ制御装置において、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間の接続形式が、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間の接続形式と異なることを特徴とするディスクアレイ制御装置。

【請求項 2】 前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセレクト部を介して接続され、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間は、それぞれセレクト部を介さずに接続されていることを特徴とする請求項 1 に記載のディスクアレイ制御装置。

【請求項 3】 前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセレクト部を介して接続され、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間はそれぞれは 1 対 1 接続されていることを特徴とする請求項 1 に記載のディスクアレイ制御装置。

【請求項 4】 前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセレクト部を介して接続され、前記複数のチャネルインターフェース部、前記複数のディスクインターフェース部、及び前記共有メモリ部は共有バスに接続されていることを特徴とする請求項 1 に記載のディスクアレイ制御装置。

【請求項 5】 前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はスイッチを用いた相互結合網によって接続され、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間はそれぞれ直接接続されていることを特徴とする請求項 1 に記載のディスクアレイ制御装置。

【請求項 6】 前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はスイッチを用いた相互結合網によって接続され、前記複数のチャネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間はそれぞれ 1 対 1 接続されていることを特徴とする請求項 1 に記載のディスクアレイ制御装置。

【請求項 7】 ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記複数のチャネルインターフェース部と前記複数のディスクインターフェース部とに接続され、前記磁気ディスク装置に対しリード／ライトされるデータを一時的に格納するキャッシュメモリ部と、前記複数のチャネルインターフェース部と前記複数のディスクインターフェース部とに接続され、前記チャネルインターフェース部及び前記ディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報を格納する前記共有メモリ部とを有する制御装置において、前記複数のチャネルインターフェース部と前記複数のディスクインターフェース部から前記キャッシュメモリ部に接続されるアクセスパスの本数は、前記複数のチャネルインターフェース部と前記複数のディスクインターフェース部から前記共有メモリ部に接続されるアクセスパスの本数より少ないことを特徴とするディスクアレイ制御装置。

【請求項 8】 セレクト部をさらに有し、前記各チャネルインターフェース部及び前記各ディスクインターフェース部と前記セレクト部とは、それぞれアクセスパスにより 1 対 1 に接続され、前記セレクト部と前記キャッシュメモリ部とは、アクセスパスにより接続され、前記セレクト部を介して前記キャッシュメモリ部と接続された前記チャネルインターフェース部及び前記ディスクインターフェース部と前記セレクト部とを接続する前記アクセスパスの本数は、前記セレクト部と前記メモリ部を接続する前記アクセスパスの本数より多く、前記各チャネルインターフェース部及び前記各ディスクインターフェース部と前記共有メモリ部との間はそれぞれアクセスパスにより 1 対 1 に接続されていることを特徴とする請求項 7 に記載のディスクアレイ制御装置。

【請求項 9】 ホストコンピュータとの複数のインターフェース部と、磁気ディスク装置との複数のインターフェース部と、前記磁気ディスク装置のデータを一時的に格納する物理的に独立したキャッシュメモリ部と、制御情報を格納する物理的に独立した共有メモリ部とを有し、前記ホストコンピュータとの各インターフェース部及び前記磁気ディスク装置との各インターフェース部は、それぞれマイクロプロセッサと、前記キャッシュメモリ部へのアクセスを制御するキャッシュメモリアクセス制御部と、前記共有メモリ部へのアクセスを制御する共有メ

10

20

30

40

50

モリアクセス制御部とを有するディスクアレイ制御装置において、前記各キャッシュメモリアクセス制御部と前記キャッシュメモリ部との間は、セレクト部を介してアクセスパスにより接続されており、前記各共有メモリアクセス制御部と前記共有メモリ部との間は、それぞれアクセスパスにより 1 対 1 接続されていることを特徴とするディスクアレイ制御装置。

【請求項 10】ホストコンピュータとの複数のインターフェース部と、磁気ディスク装置との複数のインターフェース部と、前記磁気ディスク装置のデータを一時的に格納する物理的に独立したキャッシュメモリ部と、制御情報を格納する物理的に独立した共有メモリ部とを有し、前記ホストコンピュータとの各インターフェース部及び前記磁気ディスク装置との各インターフェース部は、それぞれマイクロプロセッサと、前記キャッシュメモリ部へのアクセスを制御するキャッシュメモリアクセス制御部と、前記共有メモリ部へのアクセスを制御する共有メモリアクセス制御部とを有するディスクアレイ制御装置において、前記各キャッシュメモリアクセス制御部と前記キャッシュメモリ部との間は、スイッチを用いた相互結合網によって接続されており、前記各共有メモリアクセス制御部と前記共有メモリ部との間は、それぞれアクセスパスにより 1 対 1 接続されていることを特徴とするディスクアレイ制御装置。

【請求項 11】前記各キャッシュメモリアクセス制御部と前記キャッシュメモリ部との間の前記アクセスパスの帯域幅を、前記ホストコンピュータと前記ホストコンピュータとのインターフェース部との間の最大の全帯域幅の 2 倍以上としたことを特徴とする請求項 9 または請求項 10 の何れかに記載のディスクアレイ制御装置。

【請求項 12】前記共有メモリ部及び前記キャッシュメモリ部を複数有し、該複数の共有メモリ部及び前記複数のキャッシュメモリ部はそれぞれ二重化されていることを特徴とする請求項 1 乃至請求項 11 の何れかに記載のディスクアレイ制御装置。

【請求項 13】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースと、前記チャンネルインターフェースと前記ディスクインターフェースを 1 対 1 接続可能な第一のアクセスパスと、前記チャンネルインターフェースと前記ディスクインターフェースを接続する第二のアクセスパスを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは前記第一のアクセスパスを介して伝送され、また、データ伝送に関する制御情報は前記第二のアクセスパスを介して伝送され、さらに、前記第一のアクセスパスの帯域幅は、前記第二のアクセスパスの帯域幅よりも広いことを特徴とするディスクアレイ制御装置。

【請求項 14】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスク

インターフェースと、前記チャンネルインターフェースと前記ディスクインターフェースを 1 対 1 接続可能なデータ用ネットワークと、前記チャンネルインターフェースと前記ディスクインターフェースを接続する制御情報用ネットワークを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは前記データ用ネットワークを介して伝送され、データ伝送に関する制御情報は前記制御情報用ネットワークを介して、前記データ用ネットワークとは独立に、伝送されることを特徴とするディスクアレイ制御装置。

【請求項 15】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースを備え、前記チャンネルインターフェースは第一及び第二のアクセス制御部を備え、前記ディスクインターフェースは第三及び第四のアクセス制御部を備え、前記第一のアクセス制御部と前記第三のアクセス制御部は 1 対 1 接続可能であり、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記第一のアクセス制御部及び前記第三のアクセス制御部を介して伝送され、データ伝送に関する制御情報は、前記第二のアクセス制御部及び前記第四のアクセス制御部を介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項 16】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースと、前記複数のチャンネルインターフェースと前記複数のディスクインターフェースを接続するスイッチを備え、データ伝送に関する制御情報は、前記スイッチを介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項 17】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースと、前記複数のチャンネルインターフェースと前記複数のディスクインターフェースを接続する第一のスイッチと、前記複数のチャンネルインターフェースと前記複数のディスクインターフェースを接続する第二のスイッチを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記第一のスイッチを介して伝送され、データ伝送に関する制御情報は、前記第二のスイッチを介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項 18】ホストコンピュータとの複数のチャンネルインターフェースと、ディスク装置との複数のディスクインターフェースと、前記チャンネルインターフェースと第一のアクセスパスを介して接続され、前記ディスクインターフェースと第二のアクセスパスを介して接続されたスイッチを備え、前記チャンネルインターフェースは、データ伝送に関する制御情報を前記第一のパスを介して前記複数のディスクインターフェースと通信し、前記ディスクインターフェースは、データ伝送に関する制御情

10

20

30

40

50

報を前記第二のバスを介して前記複数のチャネルインターフェースと通信することを特徴とするディスクアレイ制御装置。

【請求項 19】ディスク装置に接続されたディスクアレイ制御装置であって、外部装置との第一のインターフェースと、外部装置との第二のインターフェースと、前記第一のインターフェースと前記第二のインターフェースを 1 対 1 接続可能な第一のアクセスバスと、前記第一のインターフェースと前記第二のインターフェースを接続する第二のアクセスバスを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記第一のアクセスバスを介して伝送され、データ伝送に関する制御情報は、前記第二のアクセスバスを介して伝送され、前記第一のアクセスバスの帯域幅は前記第二のアクセスバスの帯域幅よりも広いことを特徴とするディスクアレイ制御装置。

【請求項 20】ディスク装置に接続されたディスクアレイ制御装置であって、外部装置との第一のインターフェースと、外部装置との第二のインターフェースと、前記第一のインターフェースと前記第二のインターフェースを 1 対 1 接続可能なデータ用ネットワークと、前記第一のインターフェースと前記第二のインターフェースを接続する制御情報用ネットワークを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記データ用ネットワークを介して伝送され、データ伝送に関する制御情報は前記制御情報用ネットワークを介して、前記データ用ネットワークとは独立に、伝送されることを特徴とするディスクアレイ制御装置。

【請求項 21】ディスク装置に接続されたディスクアレイ制御装置であって、外部装置との第一のインターフェースと、外部装置との第二のインターフェースと、前記第一のインターフェースと前記第二のインターフェースを接続する第一のスイッチと、前記第一のインターフェースと前記第二のインターフェースを接続する第二のスイッチと、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは前記第一のスイッチを介して伝送され、データ伝送に関する制御情報は前記第二のスイッチを介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項 22】ディスク装置に接続されたディスクアレイ制御装置であって、外部装置との複数の第一のインターフェースと、外部装置との複数の第二のインターフェースと、前記複数の第一のインターフェースと第一のアクセスバスを介して接続され、前記複数の第二のインターフェースと第二のアクセスバスを介して接続されたスイッチを備え、前記第一のインターフェースは、データ伝送に関する制御情報を前記第一のアクセスバスを介して、前記複数の第二のインターフェースと通信し、前記第二のインターフェースは、データ伝送に関する制御情

報を前記第二のアクセスバスを介して、前記複数の第一のインターフェースと通信することを特徴とするディスクアレイ制御装置。

【請求項 23】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記チャネルインターフェース部と前記ディスクインターフェース部を 1 対 1 接続可能な第一のアクセスバスと、前記チャネルインターフェース部と前記ディスクインターフェース部を接続する第二のアクセスバスを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記第一のアクセスバスを介して伝送され、データ伝送に関する制御情報は、前記第二のアクセスバスを介して伝送され、前記第一のアクセスバスの帯域幅は前記第二のアクセスバスの帯域幅よりも広いことを特徴とするディスクアレイ制御装置。

【請求項 24】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記チャネルインターフェース部と前記ディスクインターフェース部を 1 対 1 接続可能なデータ用ネットワークと、前記チャネルインターフェース部と前記ディスクインターフェース部を接続する制御情報用ネットワークを備え、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは、前記データ用ネットワークを介して伝送され、データ伝送に関する制御情報は前記制御情報用ネットワークを介して、前記データ用ネットワークとは独立に、伝送されることを特徴とするディスクアレイ制御装置。

【請求項 25】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記チャネルインターフェース部と前記ディスクインターフェース部を接続する第一のスイッチと、前記チャネルインターフェース部と前記ディスクインターフェース部を接続する第二のスイッチと、前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータは前記第一のスイッチを介して伝送され、データ伝送に関する制御情報は前記第二のスイッチを介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項 26】ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部と、ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記複数のチャネルインターフェース部と第一のアクセスバスを介して接続され、前記複数のディスクインターフェース部と第二のアクセスバスを介して接続されたスイッチを備え、前記チャネルイ

10

20

30

40

50

ンターフェース部は、データ伝送に関する制御情報を前記第一のアクセスパスを介して、前記複数のディスクインターフェース部と通信し、前記ディスクインターフェース部は、データ伝送に関する制御情報を前記第二のアクセスパスを介して、前記複数のチャンネルインターフェース部と通信することを特徴とするディスクアレイ制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、データを複数の磁気ディスク装置に格納するディスクアレイ装置の制御装置に関する。

【0002】

【従来の技術】半導体記憶装置を記憶媒体とするコンピュータの主記憶のI/O性能に比べて、磁気ディスクを記憶媒体とするディスクサブシステム（以下「サブシステム」という。）のI/O性能は3～4桁程度小さく、従来からこの差を縮めること、すなわちサブシステムのI/O性能を向上させる努力がなされている。サブシステムのI/O性能を向上させるための1つの方法として、複数の磁気ディスク装置でサブシステムを構成し、データを複数の磁気ディスク装置に格納する、いわゆるディスクアレイと呼ばれるシステムが知られている。

【0003】図2は、従来のディスクアレイの構成を示す。ホストコンピュータ50とディスクアレイ制御装置2との間のデータ転送を実行する複数のチャンネルIF部11と、磁気ディスク装置20とディスクアレイ制御装置2間のデータ転送を実行する複数のディスクIF部12と、磁気ディスク装置20のデータを一時的に格納するキャッシュメモリ部14と、ディスクアレイ制御装置2に関する制御情報（例えば、チャンネルIF部及びディスクIF部とキャッシュメモリ部14との間のデータ転送制御に関する情報）を格納する共有メモリ部15とを備え、キャッシュメモリ部14および共有メモリ部15は全てのチャンネルIF部11及びディスクIF部12からアクセス可能な構成となっている。このディスクアレイでは、チャンネルIF部11及びディスクIF部12と共有メモリ部15との間、及び、チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14との間は1対1に接続される。以下、このような接続形態をスター接続と呼ぶ。

【0004】チャンネルIF部11は、ホストコンピュータ50と接続するためのインターフェース及びホストコンピュータ50に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスクIF部12は、磁気ディスク装置20と接続するためのインターフェース及び磁気ディスク装置20に対する入出力を制御するマイクロプロセッサ（図示せず）を有している。また、ディスクIF部12は、RAID機能の実行も行う。

【0005】図3は、他の従来のディスクアレイの構成を示す。ホストコンピュータ50とディスクアレイ制御装置3間のデータ転送を実行する複数のチャンネルIF部11と、磁気ディスク装置20とディスクアレイ制御装置3間のデータ転送を実行する複数のディスクIF部12と、磁気ディスク装置20のデータを一時的に格納するキャッシュメモリ部14と、ディスクアレイ制御装置3に関する制御情報（例えば、チャンネルIF部及びディスクIF部とキャッシュメモリ部14との間のデータ転送制御に関する情報）を格納する共有メモリ部15を備え、各チャンネルIF部11及びディスクIF部12と共有メモリ部15間は共有バス130で接続され、各チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14間は共有バス131で接続される。以下、このような接続形態を共有バス接続形式と呼ぶ。

【0006】

【発明が解決しようとする課題】ディスクアレイのアーキテクチャーをスケーラブルなものとするには、ディスク制御装置に接続するディスク容量（論理ボリューム数）に応じ、ディスクIF部を増設し、また、必要なホストコンピュータとのチャンネル数に応じて、ディスクアレイ制御装置内のチャンネルIF部を増設する必要がある。しかし、第3図に示した共有バス接続形式のディスクアレイ制御装置では、一旦実装した共有バスの転送能力をチャンネルIF部、ディスクIF部の増設に応じて変更することはできないので、チャンネルIF部、ディスクIF部の増設に柔軟に対応することが困難である。

【0007】また、第3図に示した共有バス接続形式のディスクアレイ制御装置では、ホストコンピュータとディスクアレイ制御装置との間のデータ転送を実行するチャンネルIF部に設けられたホストコンピュータに対する入出力を制御するマイクロプロセッサ、及び磁気ディスク装置とディスクアレイ制御装置との間のデータ転送を実行するディスクIF部に設けられた磁気ディスク装置に対する入出力を制御するマイクロプロセッサに高性能なプロセッサを使用した場合に、これらのプロセッサの性能に比べて、共有バスの転送能力がボトルネックになり、プロセッサの高速化に追従することが困難となる。

【0008】さらに、第3図に示した共有バス接続形式では、共有バスに接続された複数のチャンネルIF部（または複数のディスクIF部）の何れかのチャンネルIF部（またはディスクIF部）に障害が発生した場合に、障害の発生したチャンネルIF部（またはディスクIF部）を特定することが困難である。

【0009】一方、第2図に示したスター接続形式のディスクアレイ制御装置では、共有メモリ部またはキャッシュメモリ部に接続したアクセスパス数に比例して内部バス性能が増加させることができるので、チャンネルIF部、ディスクIF部の増設、または使用するプロセッサの性能に応じて、内部バス性能を増加させることが可能

である。また、チャンネル I F 部及びディスク I F 部とキャッシュメモリ部との間、チャンネル I F 部及びディスク I F 部と共有メモリ部との間がスター接続されているため、障害の発生したチャンネル I F 部（またはディスク I F 部）を特定することも容易である。

【0010】スター接続形式のディスクアレイ制御装置では、搭載されるチャンネル I F 部またはディスク I F 部の数を増やした場合、チャンネル I F 部及びディスク I F 部とキャッシュメモリ部との間、及びチャンネル I F 部及びディスク I F 部と共有メモリ部との間のアクセスパス数も増えることになる。また、ホストコンピュータとディスクアレイ制御装置との間の接続にファイバチャンネル等の高速チャンネルの採用等により、ディスクアレイ制御装置に要求されるスループットはさらに増大する方向にあり、このスループットの向上の要求を満たすためには、チャンネル I F 部及びディスク I F 部とキャッシュメモリ部との間、及びチャンネル I F 部及びディスク I F 部と共有メモリ部との間のアクセスパス数を増やし、内部パス性能を向上させることが必要となる。

【0011】しかし、キャッシュメモリに格納される 1 つのデータのデータ量は、共有メモリに格納される 1 つの制御情報のデータ量よりかなり大きい。一例を挙げれば、メインフレームに接続されるディスク制御装置では、キャッシュメモリに格納される 1 つのデータは数 K バイト程度（例えば 2 K バイト）であるのに対し、共有メモリに格納される 1 つの制御情報は数バイト程度（例えば 4 バイト）である。また、オープン系のホストコンピュータに接続されるディスク制御装置では、キャッシュメモリに格納される 1 つのデータは数十バイト程度（例えば 64 バイト）であるのに対し、共有メモリに格納される 1 つの制御情報は数バイト程度（例えば 4 バイト）である。したがって、チャンネル I F 部及びディスク I F 部とキャッシュメモリ部との間で転送されるデータ量は、チャンネル I F 部及びディスク I F 部と共有メモリ部との間で転送されるデータ量に比べ、かなり多いので、チャンネル I F 部及びディスク I F 部とキャッシュメモリ部との間のアクセスパスのデータ幅は、チャンネル I F 部及びディスク I F 部と共有メモリ部との間のアクセスパスのデータ幅より広くとる必要がある。例えば、前者のアクセスパスは、16 ビット幅のバスで構成され、後者は、4 ビット幅のバスで構成される。そのため、チャンネル I F 部及びディスク I F 部とキャッシュメモリ部との間のアクセスパスの本数を増やすと、それらのアクセスパスを接続するキャッシュメモリ部の L S I のピン数が不足するという問題が生じる。また、ディスクアレイ制御装置のホストコンピュータへの応答時間を短くするためには、共有メモリ部に格納された制御情報へのアクセス時間をできるだけ短くすることも必要である。

【0012】そこで、本発明の目的は、キャッシュメモリ及び共有メモリに格納されるデータの特性及びこれら

のメモリへのアクセス特性を考慮した、スループットの高く、ディスクアレイ制御装置、及びそれを用いたサブシステムを提供することにある。

【0013】より具体的には、本発明の目的は、チャンネル I F 部及びディスク I F 部とキャッシュメモリ部との間のアクセスパスはスループットが高く、チャンネル I F 部及びディスク I F 部と共有メモリ部との間のアクセスパスはスループットが高く、かつアクセス時間が短いディスクアレイ制御装置、及びそれを用いたサブシステムを提供することにある。

【0014】

【課題を解決するための手段】上記目的は、ホストコンピュータとのインターフェースを有する複数のチャンネルインターフェース部と、磁気ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、前記磁気ディスク装置に対しリード/ライトされるデータを一時的に格納するキャッシュメモリ部と、チャンネルインターフェース部及びディスクインターフェース部と前記キャッシュメモリ部との間のデータ転送に関する制御情報を格納する共有メモリ部とを有し、各チャンネルインターフェース部は、前記ホストコンピュータとのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行し、各ディスクインターフェース部は、前記磁気ディスク装置とのインターフェースと前記キャッシュメモリ部との間のデータ転送を実行するディスクアレイ制御装置において、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間の接続形式が、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間の接続形式と異なることを特徴とするディスクアレイ制御装置により達成される。

【0015】好ましくは、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセレクト部を介して接続し、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記共有メモリ部との間は、それぞれセレクト部を介せず直接接続する。

【0016】また、好ましくは、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はセレクト部を介して接続し、前記複数のチャンネルインターフェース部、前記複数のディスクインターフェース部、及び前記共有メモリ部は共有バスに接続する。

【0017】また、好ましくは、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部と前記キャッシュメモリ部との間はスイッチを用いた相互結合網によって接続し、前記複数のチャンネルインターフェース部及び前記複数のディスクインターフェース部



ース部と前記共有メモリ部との間はそれぞれ直接接続する。

【0018】その他、本願が開示する課題、及びその解決方法は、発明の実施形態の欄及び図面により明らかにされる。

【0019】

【発明の実施の形態】以下、本発明の実施例を図面を用いて説明する。

【0020】〔実施例1〕図1に、本発明の一実施例を示す。

【0021】ディスクアレイ制御装置1は、ホストコンピュータ50との2つのインターフェース部（チャンネルIF部）11と、磁気ディスク装置20との2つのインターフェース部（ディスクIF部）12と、2つのセクタ部13と、2つのキャッシュメモリ部14と、2つの共有メモリ部15と、アクセスパス0：135と、アクセスパス1：136と、アクセスパス2：137とを有する。

【0022】チャンネルIF部11は、ホストコンピュータ50との2つのIF（ホストIF）102と、ホストコンピュータ50に対する入出力を制御する2つのマイクロプロセッサ101と、キャッシュメモリ部14へのアクセスを制御するアクセス制御部（CMアクセス制御部）104と、共有メモリ部15へのアクセスを制御するアクセス制御部（SMアクセス制御部）105とを有し、ホストコンピュータ50とキャッシュメモリ部14間のデータ転送、及びマイクロプロセッサ101と共有メモリ部15間の制御情報の転送を実行する。マイクロプロセッサ101及びホストIF102は内部バス106によって接続され、CMアクセス制御部104は2つのホストIF102に直接接続されている。また、SMアクセス制御部105は2つのマイクロプロセッサ101に直接接続されている。

【0023】ディスクIF部12は、磁気ディスク装置20との2つのIF（ドライブIF）103と、磁気ディスク装置20に対する入出力を制御する2つのマイクロプロセッサ101と、キャッシュメモリ部14への1つのアクセス制御部（CMアクセス制御部）104と、共有メモリ部15への1つのアクセス制御部（SMアクセス制御部）105を有し、磁気ディスク装置20とキャッシュメモリ部14間のデータ転送、及びマイクロプロセッサ101と共有メモリ部15間の制御情報の転送を実行する。マイクロプロセッサ101及びドライブIF103は内部バス106によって接続され、CMアクセス制御部104は2つのドライブIF103に直接接続されている。また、SMアクセス制御部105は2つのマイクロプロセッサ101に直接接続されている。ディスクIF部はRAID機能の実行も行う。

【0024】キャッシュメモリ部14は、キャッシュメモリ（CM）コントローラ107とメモリモジュール1

09を有し、磁気ディスク装置20へ記録するデータを一時的に格納する。

【0025】共有メモリ部15は、共有メモリ（SM）コントローラ108とメモリモジュール109とを有し、ディスクアレイ制御装置1の制御情報（例えば、チャンネルIF部及びディスクIF部とキャッシュメモリ部14との間のデータ転送制御に関する情報）等を格納する。

【0026】CMアクセス制御部104には2本のアクセスパス0：135を接続し、それらを2つの異なるセクタ部13にそれぞれ接続する。セクタ部13には2本のアクセスパス1：136を接続し、それらを2つの異なるCMコントローラ107にそれぞれ接続する。したがってCMコントローラ107には、2つのセクタ部から1本ずつ、計2本のアクセスパス1：136が接続される。こうすることにより、1つのCMアクセス制御部104から1つのCMコントローラ107へのアクセスルートが2つとなる。これにより、1つのアクセスパスまたはセクタ部13に障害が発生した場合でも、もう1つのアクセスルートによりキャッシュメモリ部14へアクセスすることが可能となるため、耐障害性を向上させることができる。

【0027】SMアクセス制御部105には2本のアクセスパス2：137を接続し、そのアクセスパス2：137を2つの異なるSMコントローラにそれぞれ接続する。したがって、SMコントローラ108には、2つのチャンネルIF部11及び2つのディスクIF部12から1本ずつ、計4本のアクセスパス2：137が接続される。本実施例では、1つのSMアクセス制御部105と1つのSMコントローラ108の間には1本のアクセスパス2：137を接続したが、このアクセスパス2：137を2本に増やすことにより、1つのSMアクセス制御部105から1つのSMコントローラ108へのアクセスルートが2つとなるため、耐障害性を向上させることができる。

【0028】セクタ部13には、2つのチャンネルIF部11と、2つのディスクIF部12からそれぞれ1本ずつ、計4本のアクセスパス0：135が接続される。また、セクタ部13には、2つのキャッシュメモリ部14へのアクセスパス1：136が1本ずつ、計2本接続される。

【0029】アクセスパス0：135とアクセスパス1：136の間に上記のようなバス数の関係があるため、セクタ部13ではチャンネルIF部11及びディスクIF部12からの4本のアクセスパス0：135からの要求の内、キャッシュメモリ部14へのアクセスパス1：136の数に相当する2個だけを選択して実行する機能を持つ。

【0030】本実施例の大きな特徴は、チャンネルIF部11及びディスクIF部12とキャッシュメモリ部14



との間の接続形式と、チャンネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 5 との間の接続形式とが異なる点にある。このような構成にした理由を図 1 及び図 2 を用いて説明する。図 2 に示したスター接続形式のディスクアレイ制御装置 2 では、ディスクアレイ制御装置 2 に搭載されるチャンネル I F 部 1 1 またはディスク I F 部 1 2 の数を増やした場合、チャンネル I F 部 1 1 及びディスク I F 部 1 2 とキャッシュメモリ部 1 4 との間、及びチャンネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 5 との間のアクセスパス数も増えることになる。また、スター接続形式のディスクアレイ制御装置 2 において、スループットを向上させるためには、チャンネル I F 部 1 1 及びディスク I F 部 1 2 とキャッシュメモリ部 1 4 との間、及びチャンネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 5 との間のアクセスパス数を増やし、内部バス性能を向上させることが有効である。

【0031】しかし、キャッシュメモリ部 1 4 に格納される 1 つのデータのデータ量は、共有メモリ部 1 5 に格納される 1 つの制御情報のデータ量よりかなり大きい。一例を挙げれば、メインフレームに接続されるディスク制御装置では、キャッシュメモリ部 1 4 に格納される 1 つのデータは数 K バイト程度（例えば 2 K バイト）であるのに対し、共有メモリ部 1 5 に格納される 1 つの制御情報は数バイト程度（例えば 4 バイト）である。また、オープン系のホストコンピュータに接続されるディスク制御装置では、キャッシュメモリ部 1 4 に格納される 1 つのデータは数十バイト程度（例えば 64 バイト）であるのに対し、共有メモリ部 1 5 に格納される 1 つの制御情報は数バイト程度（例えば 4 バイト）である。したがって、チャンネル I F 部 1 1 及びディスク I F 部 1 2 とキャッシュメモリ部 1 4 との間で転送されるデータ量は、チャンネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 5 との間で転送されるデータ量に比べかなり多いので、チャンネル I F 部 1 1 及びディスク I F 部 1 2 とキャッシュメモリ部 1 4 との間のアクセスパスのデータ幅は、チャンネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 5 との間のアクセスパスのデータ幅より広くとる必要がある。例えば、前者のアクセスパスは 16 ビット幅のバスで構成され、後者は、8 ビット幅のバスで構成される。そのため、チャンネル I F 部 1 1 及びディスク I F 部 1 2 とキャッシュメモリ部 1 4 との間のアクセスパス本数を増やすと、キャッシュメモリ部 1 4 内のキャッシュメモリコントローラ（図 2 ではキャッシュメモリコントローラを図示していない）の L S I のピン数不足、またはキャッシュメモリ部 1 4 を実装するパッケージにおいてコネクタのピン数不足という問題が生じる。そこで、本実施例では、図 1 に示すように、チャンネル I F 部 1 1 及びディスク I F 部 1 2 とキャッシュメモリ部 1 4 との間をセレクト部 1 3 を介して接続すること

により、キャッシュメモリ部 1 4 に直接接続されるアクセスパス数を削減している。

【0032】一方、上述したように、共有メモリ部 1 3 へ格納する 1 つの制御情報のデータ長はキャッシュメモリ部 1 4 に格納する 1 つのデータのデータ長に比べかなり小さいので、チャンネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 5 との間のアクセスパスのデータ幅は、チャンネル I F 部 1 1 及びディスク I F 部 1 2 とキャッシュメモリ部 1 4 との間のアクセスパスのデータ幅の半分以上とすることが可能である。したがって、共有メモリ部 1 5 へのアクセスパス数を増やしても共有メモリ部内の共有メモリメモリコントローラ（図 2 では共有メモリコントローラを図示していない）の L S I のピン数不足等の問題が生じることは少ない。

【0033】また、ディスクアレイ制御装置 1 のホストコンピュータ 50 への応答時間を短くするためには、共有メモリ部 1 5 に格納される制御情報へのアクセス時間をできるだけ短くする必要もある。しかし、図 1 に示した CM アクセス制御部 104 と CM コントローラ 107 間のように、SM アクセス制御部 105 と SM コントローラ 108 との間をセレクト部を介して接続すると、セレクト部での処理のオーバーヘッドにより、共有メモリ部 1 5 に格納される制御情報へのアクセス時間を短くすることができない。

【0034】そこで、本実施例では、チャンネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 5 との間をセレクト部を介さず直接接続することにより、チャンネル I F 部 1 1 及びディスク I F 部 1 2 と共有メモリ部 1 5 間に複数のアクセスパスを設け、スループットを向上させる一方、セレクト部での処理オーバーヘッドをなくし、チャンネル I F 部 1 1、及びディスク I F 部 1 2 から共有メモリ部 1 5 へのアクセス時間を短縮にしている。

【0035】なお、本実施例では、耐障害性の向上という観点から、セレクト部 13、キャッシュメモリ部 14、及び共有メモリ部 15 をそれぞれ二重化しているが、これらを二重化しなくても、上述の効果をえられることは言うまでもない。

【0036】図 4 は、CM アクセス制御部 104 内の構成を示している。CM アクセス制御部 104 は、セレクト 302 と、アドレス、コマンド、データを一時格納するバケットバッファ 303 と、セレクト部 13 に繋がるアクセスパス 0 : 135 とのバス I F 301 と、データのエラーチェック部 300 と、データ転送制御部 310 を有する。セレクト 302 の 2 つのポートはデータ線 210 でホスト I F 102 あるいはドライブ I F 103 に接続される。また、セレクト 302 の他の 2 つのポートはバス I F 301 に接続される。バス I F 301 はアクセスパス 0 : 135 でセレクト部 13 に接続される。データ転送制御部 310 は、制御線 1 : 211 でホスト I F 102 あるいはドライブ I F 103 に接続され、制御

15

線 2 : 2 1 2 でセクタ部 1 3 内のデータ転送制御部 3 1 5 に接続される。また、データ転送制御部 3 1 0 は、アービタ 3 0 8 によりホスト I F 1 0 2 あるいはドライブ I F 1 0 3 からのアクセス要求のアービトレーションを行い、セクタ 3 0 2 の切り替えを行う。

【0037】図 6 は、セクタ部 1 3 内の構成を示している。セクタ部 1 3 は、チャンネル I F 部 1 1 及びディスク I F 部 1 2 に繋がるアクセスパス 0 : 1 3 5 との 4 つのバス I F 3 0 1 と、CM コントローラ 1 0 7 に繋がるアクセスパス 1 : 1 3 6 との 2 つのバス I F 3 0 1 と、両者間を互いに接続するセクタ 3 0 6 と、パケットバッファ 3 0 3 と、データのエラーチェック部 3 0 0 と、CM アクセス制御部 1 0 4 から送出されたアドレス及びコマンドを解析するアドレス・コマンド ( a d r 、 c m d ) 解析部 3 0 5 と、データ転送制御部 3 1 5 を有する。データ転送制御部 3 1 5 は、制御線 2 : 2 1 2 で CM アクセス制御部 1 0 4 内のデータ転送制御部 3 1 0 に接続され、制御線 3 : 2 1 3 で CM コントローラ 1 0 7 内のデータ転送制御部 3 1 5 に接続される。また、データ転送制御部 3 1 5 は、アービタ 3 0 8 により、a d r 、 c m d 解析部 3 0 5 で解析した 4 本のアクセスパス 0 : 1 3 5 からのアクセス要求のアービトレーションを行い、セクタ 3 0 6 の切り替えを行う。パケットバッファ 3 0 3 は、アクセスパス 0 : 1 3 5 側のバスとアクセスパス 1 : 1 3 6 側のバスでデータ転送速度に差がある場合、速度差を吸収するために、転送するデータの一部または全部をバッファリングする。

【0038】a d r 、 c m d 解析部 3 0 5 は、アドレス及びコマンドを格納するバッファと、a d r 抽出部と、c m d 抽出部を有する ( 図示していない ) 。 a d r 、 c m d 解析部 3 0 5 では、CM アクセス制御部 1 0 4 に接続される 4 本のアクセスパス 0 : 1 3 5 それぞれに 1 つずつ割り当てられたバッファに、アドレス、コマンドを格納する。a d r 抽出部及び c m d 抽出部では、アクセスする CM コントローラ 1 0 7 とアクセスの種類を割り出し、データ転送制御部 3 1 5 内のアービタ 3 0 8 へ送出する。

【0039】図 7 は、キャッシュメモリ部 1 4 内の構成を示している。キャッシュメモリ部 1 4 は、CM コントローラ 1 0 7 とメモリモジュール 1 0 9 を有する。CM コントローラ 1 0 7 は、セクタ部 1 3 に繋がるアクセスパス 1 : 1 3 6 との 2 つのバス I F 3 0 1 と、セクタ 3 0 4 と、データを一時格納するパケットバッファ 3 0 3 と、データのエラーチェック部 3 0 0 と、メモリモジュール 1 0 9 へのアクセスを制御するメモリ制御部 3 0 7 と、CM アクセス制御部 1 0 4 から送出されたアドレス及びコマンドを解析する a d r 、 c m d 解析部 3 0 5 と、データ転送制御部 3 1 5 を有する。データ転送制御部 3 1 5 は、制御線 3 : 2 1 3 でセクタ部 1 3 内のデータ転送制御部 3 1 5 に接続される。また、データ転

16

送制御部 3 1 5 は、アービタ 3 0 8 により、a d r 、 c m d 解析部 3 0 5 で解析した 2 本のアクセスパス 1 : 1 3 6 からのアクセス要求のアービトレーションを行い、セクタ 3 0 4 の切り替えを行う。

【0040】a d r 、 c m d 解析部 3 0 5 は、バッファと、a d r 抽出部と、c m d 抽出部を有する ( 図示していない ) 。 a d r 、 c m d 解析部 3 0 5 では、CM コントローラ 1 0 7 に接続される 2 本のアクセスパス 1 : 1 3 6 それぞれに 1 つずつ割り当てられたバッファに、アドレス、コマンドを格納する。a d r 抽出部及び c m d 抽出部では、アクセスするメモリのアドレスとアクセスの種類を割り出し、メモリ制御部 3 0 7 へ送出する。また、2 本のアクセスパス 1 : 1 3 6 からのアクセス要求をデータ転送制御部 3 1 5 内のアービタ 3 0 8 へ送出する。

【0041】次に、キャッシュメモリ部 1 4 へのアクセス時の手順について述べる。キャッシュメモリ部 1 4 へアクセスする場合、マイクロプロセッサ 1 0 1 は、ホスト I F 1 0 2 あるいはドライブ I F 1 0 3 へ、キャッシュメモリ部 1 4 へのアクセス開始を指示する。

【0042】アクセス開始の指示を受けたホスト I F 1 0 2 あるいはドライブ I F 1 0 3 は、制御線 1 : 2 1 1 により CM アクセス制御部 1 0 4 内のデータ転送制御部 3 1 0 へアクセス開始を示す信号を送出する。それとともに、データ線 2 1 0 を通してアドレス、コマンド、データ ( データの書き込み時のみ ) を送出する。

【0043】CM アクセス制御部 1 0 4 は、データ線 2 1 0 を通して送られてきたアドレス、コマンド、データ ( データの書き込み時のみ ) をパケットバッファ 3 0 3 に格納する。データ転送制御部 3 1 0 はアービトレーションを行ってバス I F 3 0 1 の使用权を決定し、セクタ 3 0 2 を切り替える。

【0044】図 9 は、キャッシュメモリ部 1 4 へデータを書き込む場合の、CM アクセス制御部 1 0 4 から CM コントローラ 1 0 7 へのアクセスの流れを示している。CM アクセス制御部 1 0 4 内のデータ転送制御部 3 1 0 は、アービトレーションによってアクセスパス 0 : 1 3 5 の使用权が決定されると、制御線 2 : 2 1 2 によってセクタ部 1 3 内のデータ転送制御部 3 1 5 へアクセス開始を示す信号 ( R E Q ) を出す ( ステップ 5 0 1 ) 。続いて、アドレス及びコマンドを送出する ( ステップ 5 0 2 ) 。

【0045】セクタ部 1 3 内のデータ転送制御部 3 1 5 は、CM アクセス制御部 1 0 4 から R E Q 信号を受け取ると、次にアクセスパス 0 : 1 3 5 を通して送られてくるアドレス及びコマンドを受信し、a d r 、 c m d 解析部 3 0 5 で解析したアクセス要求に基づいてアービトレーションを行う ( ステップ 5 0 3 ) 。アービトレーションの結果、アクセスパス 1 : 1 3 6 への接続権を得たら、データ転送制御部 3 1 5 はセクタ 3 0 6 を切り替

える(ステップ504)とともに、制御線2:212により、CMアクセス制御部104内のデータ転送制御部310へ、アクセスパス1:136への接続権が得られたことを示す信号(ACK)を返す(ステップ505)。次にデータ転送制御部315は、制御線3:213によってCMコントローラ107内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ506)。続いて、アドレス及びコマンドを送出する(ステップ507)。

【0046】CMアクセス制御部104はACK信号を受けると、パケットバッファ303からデータを読み出し、セクタ302、パスIF301を介してアクセスパス0:135へ送出する。セクタ部13は、アクセスパス0:135を通して送られてきたデータを、パスIF301及びセクタ306を介してアクセスパス1:136へ送出する(ステップ509)。

【0047】CMコントローラ107内のデータ転送制御部315は、制御線3:213によってREQ信号を受け取ると、次にアクセスパス1:136を通して送られてくるアドレス及びコマンドを受信し、addr、cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行い(ステップ508)、セクタ304を切り替える。アクセスパス1:136を通して送られてくるデータはパケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ510)。次に、パケットバッファ303からデータを読み出し、セクタ304を介してメモリモジュール109へ書き込む(ステップ511)。

【0048】メモリモジュール109へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ512)。次に、ステータスをセクタ部13を介してCMアクセス制御部104へ送出する(ステップ513)。セクタ部13内のデータ転送制御部315はステータスを受け取ると、CMコントローラ107へのREQ信号をオフする(ステップ514)。CMアクセス制御部104内のデータ転送制御部310はステータスを受け取ると、セクタ部13へのREQ信号をオフする(ステップ515)。セクタ部13内のデータ転送制御部315はCMアクセス制御部104からのREQ信号のオフを確認すると、CMアクセス制御部104へのACK信号をオフする(ステップ516)。

【0049】CMアクセス制御部104内のデータ転送制御部310はステータスを受け取ると、制御線1:211により、ホストIF102あるいはドライブIF103へキャッシュメモリ部14へのアクセスの終了を報告する。

【0050】キャッシュメモリ部14からデータを読み出す場合、CMアクセス制御部104からCMコントローラ107へのアクセスの流れは、ステップ501から508までとステップ512以降は、データの書き込みの場合と同じである。

【0051】ここでCMアクセス制御部104は、ステップ505でACK信号を受けると、データの受信待ち状態に入る。

【0052】ステップ508でメモリアクセス権を得ると、CMコントローラ107はメモリモジュール109からデータを読み出し、セクタ304、パスIF301を介してアクセスパス1:136にデータを送出する。

【0053】セクタ部13は、アクセスパス1:136を通してデータを受信すると、パスIF301及びセクタ306を介してアクセスパス0:135にデータを送出する。

【0054】CMアクセス制御部104は、アクセスパス0:135を通してデータを受信すると、セクタ302、データ線210を介してホストIF102あるいはドライブIF103へデータを送出する。

【0055】図5は、SMアクセス制御部105内の構成を示している。SMアクセス制御部104は、セクタ302と、アドレス、コマンド、データを一時格納するパケットバッファ303と、SMコントローラ108に繋がるアクセスパス2:137とのパスIF301と、データのエラーチェック部300と、データ転送制御部310を有する。セクタ302の2つのポートはデータ線220でマイクロプロセッサ101に接続される。また、セクタ302の他の2つのポートはパスIF301に接続される。パスIF301はアクセスパス2:137でSMコントローラ108に接続される。データ転送制御部310は、制御線5:221でマイクロプロセッサ101に接続され、制御線6:222でSMコントローラ108内のデータ転送制御部315に接続される。また、データ転送制御部310は、アービタ308によりマイクロプロセッサ101からのアクセス要求のアービトレーションを行い、セクタ302の切り替えを行う。

【0056】図8は、共有メモリ部15内の構成を示している。共有メモリ部15は、SMコントローラ108とメモリモジュール109を有する。SMコントローラ108は、SMアクセス制御部105に繋がるアクセスパス2:137との4つのパスIF301と、セクタ309と、データを一時格納するパケットバッファ303と、データのエラーチェック部300と、メモリモジュール109へのアクセスを制御するメモリ制御部307と、SMアクセス制御部105から送出されたアドレス及びコマンドを解析するaddr、cmd解析部305と、データ転送制御部315を有する。データ転送制御

10

20

30

40

50

部315は、制御線6:222でSMアクセス制御部105内のデータ転送制御部310に接続される。また、データ転送制御部315は、アービタ308により、adr、cmd解析部305で解析した4本のアクセスパス2:137からのアクセス要求のアービトレーションを行い、セクタ309の切り替えを行う。

【0057】adr、cmd解析部305は、バッファと、adr抽出部と、cmd抽出部を有する(図示していない)。adr、cmd解析部305では、SMコントローラ108に接続される4本のアクセスパス2:137それぞれに1つつ割り当てられたバッファに、アドレス、コマンドを格納する。adr抽出部及びcmd抽出部では、アクセスするメモリのアドレスとアクセスの種類を割り出し、メモリ制御部307へ送出する。また、4本のアクセスパス2:137からのアクセス要求をデータ転送制御部315内のアービタ308へ送出する。

【0058】次に、共有メモリ部15へのアクセス時の手順について述べる。共有メモリ部15へアクセスする場合、マイクロプロセッサ101は、制御線5:221によりSMアクセス制御部105内のデータ転送制御部310へアクセス開始を示す信号を送出する。それとともに、データ線220を通してアドレス、コマンド、データ(データの書き込み時のみ)を送出する。

【0059】SMアクセス制御部105は、データ線220を通して送られてきたアドレス、コマンド、データ(データの書き込み時のみ)をパケットバッファ303に格納する。データ転送制御部310はアービトレーションを行ってパシフ301の使用権を決定し、セクタ302を切り替える。

【0060】図10は、共有メモリ部15へデータを書き込む場合の、SMアクセス制御部105からSMコントローラ108へのアクセスの流れを示している。SMアクセス制御部105内のデータ転送制御部310は、アービトレーションによってアクセスパス2:137の使用権が決定されると、制御線6:222によってSMコントローラ108へアクセス開始を示す信号(REQ)を出す(ステップ601)。続いて、アドレス、コマンド、及びデータを連続して送出する(ステップ602)。

【0061】SMコントローラ108内のデータ転送制御部315は、制御線6:222によってREQ信号を受け取ると、次にアクセスパス2:137を通して送られてくるアドレス、コマンド、及びデータを受信する。アドレスとコマンドは、adr、cmd解析部305で解析し、アクセス要求に基づいてアービトレーションを行い(ステップ603)、セクタ309を切り替える。データはパケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部3

07へ送出し、メモリアクセスのための前処理を行う(ステップ604)。次に、パケットバッファ303からデータを読み出し、セクタ309を介してメモリモジュール109へ書き込む(ステップ605)。

【0062】メモリモジュール109へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ606)。次に、ステータスをSMアクセス制御部105へ送出する(ステップ607)。SMアクセス制御部105内のデータ転送制御部310はステータスを受け取ると、SMコントローラ108へのREQ信号をオフする(ステップ608)。

【0063】SMアクセス制御部105内のデータ転送制御部310はステータスを受け取ると、制御線5:221により、マイクロプロセッサ101へ共有メモリ部15へのアクセスの終了を報告する。

【0064】共有メモリ部15からデータを読み出す場合のSMアクセス制御部105からSMコントローラ108へのアクセスの流れは、ステップ601から604までとステップ606以降は、データの書き込みの場合と同じである。

【0065】ステップ604でメモリアクセスの前処理を行った後、SMコントローラ108はメモリモジュール109からデータを読み出し、セクタ309、パシフ301を介してアクセスパス2:137にデータを送出する。

【0066】SMアクセス制御部105は、アクセスパス2:137を通してデータを受信すると、セクタ302、データ線220を介してマイクロプロセッサ101へデータを送出する。

【0067】本実施例のディスクアレイ制御装置1では、ホストコンピュータ50とのチャンネルを2つ有するチャンネルIF部12を複数搭載しており、それらのチャンネルをそれぞれ異なるホストコンピュータに接続することが可能である。そうした場合、ホストコンピュータに接続した各チャンネルからの要求全てを並列に処理する必要がある。

【0068】ところでディスクアレイ制御装置1では、ホストコンピュータ50へデータを読み出す場合、磁気ディスク装置20に格納されたデータをディスクIF部12を介してキャッシュメモリ部14へ書き込み、そのデータをキャッシュメモリ部14から読み出して、チャンネルIF部11を介してホストコンピュータ50へ送る。またホストコンピュータ50からディスクアレイ制御装置1へデータを書き込む場合は、ホストコンピュータ50からチャンネルIF部12へ送られてきたデータをキャッシュメモリ部14へ書き込み、そのデータをキャッシュメモリ部14から読み出して、ディスクIF部12を介して磁気ディスク装置20に書き込む。さらにデ

ータのバリティを生成して磁気ディスク装置 20 に書き込むため、ディスク I/F 部 12 とキャッシュメモリ部 14 間でさらに 2~3 回のアクセスが行われる。

【0069】したがって、ホストコンピュータに接続した各チャンネルからの要求全てを並列に処理するためには、チャンネル I/F 部 11 及びディスク I/F 部 12 とキャッシュメモリ部 14 間のスループットをホストコンピュータ 50 とチャンネル I/F 部 11 間の最大のスループットの 2 倍以上にしなければならない。

【0070】本実施例では、CM アクセス制御部 104 と CM コントローラ 107 間のアクセスパスの帯域幅、及び CM コントローラ 107 とメモリモジュール 109 間の全帯域幅を、チャンネル I/F 部 11 とホストコンピュータ 50 間の最大帯域幅の 2 倍以上に設定する。これにより、全チャンネル I/F 部 11 を並列に動作させることが可能となる。

【0071】本実施例によれば、チャンネル I/F 部 11 及びディスク I/F 部 12 とキャッシュメモリ部 14 間のスループットを高くすること、チャンネル I/F 部 11 及びディスク I/F 部 12 と共有メモリ部 15 間のスループットを高く、且つアクセス時間を短くすることの両方が可能となる。これによって、スループットが高く、且つ応答時間の短いディスクアレイ制御装置を提供できる。

【0072】ここで、図 15 に示すように、チャンネル I/F 部 11 及びディスク I/F 部 12 とキャッシュメモリ部 14 間をスイッチ (SW) 16 を用いた相互結合網 140 で接続する。この場合にも、図 1 に示したセクタ部 13 を介して接続した構成と同様に、キャッシュメモリ部 14 へ複数のアクセスパスを設けることができるため、スループットを高めることが可能となる。

【0073】また図 16 に示すように、1 つの CM アクセス制御部 104 へ接続されるアクセスパス 0:135 の本数を図 1 の構成の倍の 4 本に増やしたディスクアレイ制御装置 1 においても、本実施例を実施する上で問題はない。ホスト I/F 及びドライブ I/F として、今後はファイバチャネル等のスループットが 100MB/s 以上の高速 I/F が使用されることが多くなると考えられる。チャンネル I/F 部 11 及びディスク I/F 部 12 でのスループットのバランスを考えると、1 つの CM アクセス制御部 104 に繋がる全アクセスパス 0:135 のスループットは、チャンネル I/F 部 11 内の全ホスト I/F 102、またはディスク I/F 部 12 内の全ドライブ I/F 103 のスループットと同等以上にすることが必要である。上記のようにファイバチャネル等の高速 I/F を使用する場合は、図 16 に示すように、1 つの CM アクセス制御部 104 に繋がるアクセスパス 0:135 の本数を増やすことで、アクセスパス 0:135 のスループットをホスト I/F 102 あるいはドライブ I/F 103 のスループット以上にすることができる。

【0074】なお、図 15 及び図 16 のディスクアレイ

制御装置では、耐障害性の向上という観点から、セクタ部 13、キャッシュメモリ部 14、及び共有メモリ部 15 をそれぞれ二重化しているが、これらを二重化しなくても、上述の効果をえられることは言うまでもない。

【0075】〔実施例 2〕図 1 において、2 つのキャッシュメモリ部 14 間で、メモリ領域の全部または、一部を二重化し、キャッシュメモリ部 14 へのデータの書き込み時に二重化した 2 つの領域に同じデータを書き込むことにより、データの信頼性を上げることが可能となる。

【0076】二重化した 2 つのキャッシュメモリ部 14 へデータを書き込む場合の手順は以下になる。

【0077】アクセス開始の指示を受けたホスト I/F 102 あるいはドライブ I/F 103 は、制御線 1:211 により CM アクセス制御部 104 内のデータ転送制御部 310 へアクセス開始を示す信号を送出する。それとともに、データ線 210 を通して 2 つのアドレス、コマンドと、1 つのデータを送出する。

【0078】CM アクセス制御部 104 は、データ線 210 を通して送られてきた 2 つのアドレス、コマンドと、1 つのデータをパケットバッファ 303 に格納する。データ転送制御部 310 はアービトレーションを行ってバス I/F 301 の使用权を決定し、セクタ 302 を切り替える。

【0079】図 11 は、二重化した 2 つのキャッシュメモリ部 14 へデータを書き込む場合の、CM アクセス制御部 104 から 2 つの CM コントローラ a、b:107 へのアクセスの流れを示している。CM アクセス制御部 104 内のデータ転送制御部 310 は、アービトレーションによってアクセスパス 0:135 の使用权が決定されると、制御線 2:212 によってセクタ部 13 内のデータ転送制御部 315 へアクセス開始を示す信号 (REQ) を出す (ステップ 701)。続いて、アドレス及びコマンドを 2 つ連続して送出的 (ステップ 702)。

【0080】セクタ部 13 内のデータ転送制御部 315 は、CM アクセス制御部 104 から REQ 信号を受け取ると、次にアクセスパス 0:135 を通して送られてくるアドレス及びコマンドを受信し、addr、cmd 解析部 305 で解析したアクセス要求に基づいてアービトレーションを行う (ステップ 703)。アービトレーションの結果、CM コントローラ a、b:107 への 2 本のアクセスパス 1:136 への接続権の両方を得たら、データ転送制御部 315 はセクタ 306 を切り替える (ステップ 704) とともに、制御線 2:212 により、CM アクセス制御部 104 内のデータ転送制御部 310 へ、CM コントローラ a、b:107 への接続権が得られたことを示す信号 (ACK) を返す (ステップ 705)。次にデータ転送制御部 315 は、制御線 3:213 によって 2 つの CM コントローラ 107 内のデータ

転送制御部 315 へアクセス開始を示す信号 (REQ) を出す (ステップ 706a、706b)。続いて、CM コントローラ a、b : 107 へアドレス及びコマンドを 1 つずつ送出する (ステップ 707a、707b)。

【0081】CM アクセス制御部 104 は ACK 信号を受けると、パケットバッファ 303 からデータを読み出し、セクタ 302、パス IF 301 を介してアクセスパス 0 : 135 へ送出する (ステップ 709)。セクタ部 13 は、アクセスパス 0 : 135 を通して送られてきた 1 つのデータを、パス IF 301 及びセクタ 306 を介して 2 つのアクセスパス 1 : 136 の両方へ送出する (ステップ 709a、709b)。

【0082】CM コントローラ a、b : 107 内のデータ転送制御部 315 は、制御線 3 : 213 によって REQ 信号を受け取ると、次にアクセスパス 1 : 136 を通して送られてくるアドレス及びコマンドを受信し、addr、cmd 解析部 305 で解析したアクセス要求に基づいてアービトレーションを行い (ステップ 708a、708b)、セクタ 304 を切り替える。アクセスパス 1 : 136 を通して送られてくるデータはパケットバッファ 303 に格納する。アービトレーションの結果、メモリモジュール 109 へのアクセス権を得たら、メモリの制御情報をメモリ制御部 307 へ送出し、メモリアクセスのための前処理を行う (ステップ 710a、710b)。次に、パケットバッファ 303 からデータを読み出し、セクタ 304 を介してメモリモジュール 109 へ書き込む (ステップ 711a、711b)。

【0083】メモリモジュール 109 へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部 315 においてアクセス状況を示すステータス (STATUS) を生成する (ステップ 712a、712b)。次に、ステータスをセクタ部 13 を介して CM アクセス制御部 104 へ送出する (ステップ 713a、713b)。セクタ部 13 内のデータ転送制御部 315 はステータスを受け取ると、CM コントローラ a、b : 107 への REQ 信号をそれぞれオフする (ステップ 714a、714b)。また、セクタ部 13 は CM コントローラ a、b : 107 の両方からステータスを受け取ったら、それらを続けて CM アクセス制御部へ送出する (ステップ 713)。CM アクセス制御部 104 内のデータ転送制御部 310 は 2 つのステータスを受け取ると、セクタ部 13 への REQ 信号をオフする (ステップ 715)。セクタ部 13 内のデータ転送制御部 315 は CM アクセス制御部 104 からの REQ 信号のオフを確認すると、CM アクセス制御部 104 への ACK 信号をオフする (ステップ 716)。

【0084】CM アクセス制御部 104 内のデータ転送制御部 310 はステータスを受け取ると、制御線 1 : 211 により、ホスト IF 102 あるいはドライブ IF 103 へキャッシュメモリ部 14 へのアクセスの終了を報

告する。

【0085】上記のように二重化したキャッシュメモリ部 14 にデータを二重に書き込む場合、二重に書き込むデータの一方のデータの書き込み処理を行っている間に、もう一方のデータが他のアクセスパスからの書き込み要求によって書き換えられるのを防ぐ必要がある。本実施例では、共有メモリ部 15 にキャッシュメモリ部 14 のディレクトリを格納し、キャッシュメモリ部 14 にアクセスする前に必ず、共有メモリ部 15 に格納したディレクトリにアクセス中を示すビットを立てる。これにより、キャッシュメモリ部 14 内の同じアドレスには同時に 1 つのアクセス要求しか発行されないため、二重に書き込むデータの一方のデータの書き込み処理を行っている間に、もう一方のデータが他のアクセスパスからの書き込み要求によって書き換えられるのを防ぐことができる。

【0086】〔実施例 3〕ディスクアレイ制御装置 1 では、キャッシュメモリ部 14 を複数設けた場合、あるキャッシュメモリ部 14 から別のキャッシュメモリ部 14 へデータをコピーする機能が要求される。この機能は、以下に述べる手順で実現できる。

【0087】アクセス開始の指示を受けたホスト IF 102 あるいはドライブ IF 103 は、制御線 1 : 211 により CM アクセス制御部 104 内のデータ転送制御部 310 へアクセス開始を示す信号を送出する。それとともに、データ線 210 を通して 2 つのアドレス、コマンドを送出する。2 つのうちの 1 つのアドレス及びコマンドは、コピー元のアドレスとリードコマンドで、もう 1 つのアドレス及びコマンドはコピー先のアドレスとライトコマンドである。ここでは、CM コントローラ a : 107 をコピー元、CM コントローラ b をコピー先として説明する。

【0088】CM アクセス制御部 104 は、データ線 210 を通して送られてきた 2 つのアドレス、コマンドをパケットバッファ 303 に格納する。データ転送制御部 310 はアービトレーションを行ってパス IF 301 の使用権を決定し、セクタ 302 を切り替える。

【0089】図 12 は、2 つのキャッシュメモリ部間でデータをコピーする場合の、CM アクセス制御部 104 から CM コントローラ a、b : 107 へのアクセスの流れを示している。CM アクセス制御部 104 内のデータ転送制御部 310 は、アービトレーションによってアクセスパス 0 : 135 の使用権が決定されると、制御線 2 : 212 によってセクタ部 13 内のデータ転送制御部 315 へアクセス開始を示す信号 (REQ) を出す (ステップ 801)。続いて、アドレス及びコマンドを 2 つ連続して送出する (ステップ 802)。

【0090】セクタ部 13 内のデータ転送制御部 315 は、CM アクセス制御部 104 から REQ 信号を受け取ると、次にアクセスパス 0 : 135 を通して送られて

くるアドレス及びコマンドを受信し、*adr*、*cmd*解析部305で解析したアクセス要求に基づいてアービトレーションを行う(ステップ803)。アービトレーションの結果、CMコントローラ*a*、*b*:107への2つのアクセスパス1:136への接続権の両方を得たら、データ転送制御部315はセクタ306を切り替える(ステップ804)とともに、制御線2:212により、CMアクセス制御部104内のデータ転送制御部310へ、CMコントローラ*a*、*b*:107両方への接続権が得られたことを示す信号(ACK)を返す(ステップ805)。次にデータ転送制御部315は、制御線3:213によってCMコントローラ*a*、*b*:107内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ806*a*、806*b*)。続いて、CMコントローラ*a*、*b*:107へそれぞれのアドレス及びコマンドを送出する(ステップ807*a*、807*b*)。

【0091】CMアクセス制御部104はACK信号を受けると、アクセスの終了を知らせるステータスの受信待ち状態に入る。

【0092】コピー元のCMコントローラ*a*:107内のデータ転送制御部315は、制御線3:213によってREQ信号を受け取ると、次にアクセスパス1:136を通して送られてくるアドレス及びコマンドを受信し、*adr*、*cmd*解析部305で解析したアクセス要求に基づいてアービトレーションを行い(ステップ808)、セクタ304を切り替える。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ809)。次に、メモリモジュール109からデータを読み出し(ステップ810)、セクタ304を介してアクセスパス1:136へ送出する(ステップ811*a*)。

【0093】セクタ部13は、アクセスパス1:136を通してCMコントローラ*a*:107から送られてきたデータをCMコントローラ*b*に繋がるアクセスパス1:136へ送出する。(ステップ811*b*)。

【0094】CMコントローラ*b*:107内のデータ転送制御部315は、制御線3:213によってREQ信号を受け取ると、次にアクセスパス1:136を通して送られてくるアドレス及びコマンドを受信し、*adr*、*cmd*解析部305で解析する。その後、データを受信し始めるのを待って、メモリアクセスのアービトレーションに参加する(ステップ812)。アクセスパス1:136を通して送られてくるデータはパケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ813)。次に、パケ

ットバッファ303からデータを読み出し、セクタ304を介してメモリモジュール109へ書き込む(ステップ814)。

【0095】CMコントローラ*a*、*b*:107は、それぞれCMメモリモジュール109へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ815、818)。次に、ステータスをセクタ部13へ送出する(ステップ816、819)。

【0096】セクタ部13内のデータ転送制御部315はステータスを受け取ると、CMコントローラ*a*、*b*:107へのREQ信号をそれぞれオフする(ステップ817、821)。また、セクタ部13は2つのCMコントローラ*a*、*b*:107の両方からステータスを受け取ったら、それらを続けてCMアクセス制御部へ送出する(ステップ820)。CMアクセス制御部104内のデータ転送制御部310は2つのステータスを受け取ると、セクタ部13へのREQ信号をオフする(ステップ822)。セクタ部13内のデータ転送制御部315はCMアクセス制御部104からのREQ信号のオフを確認すると、CMアクセス制御部104へのACK信号をオフする(ステップ823)。

【0097】CMアクセス制御部104内のデータ転送制御部310はステータスを受け取ると、制御線1:211により、ホストIF102あるいはドライブIF103へキャッシュメモリ部14へのアクセスの終了を報告する。

【0098】ディスクアレイ制御装置1ではまた、1つのキャッシュメモリ部14内のあるアドレスから別のアドレスへデータをコピーする機能も要求される。

【0099】この機能は、図9で示したデータの書き込み時の手順において、ステップ511のメモリモジュール109へのライトアクセスの代わりに、メモリモジュール109からデータを読み出してCMコントローラ107内のパケットバッファ303に格納し、続けてそのデータをメモリモジュール109へ書き込むという処理を行うことによって実現できる。

【0100】[実施例4]図1において、2つの共有メモリ部15間で、メモリ領域の全部または、一部を二重化し、共有メモリ部15へのデータの書き込み時に二重化した2つの領域に同じデータを書き込むことにより、データの信頼性を上げることが可能となる。

【0101】二重化した2つの共有メモリ部15へデータを書き込む場合の手順は、以下のようになる。

【0102】マイクロプロセッサ101は、制御線5:221によりSMアクセス制御部105内のデータ転送制御部310へアクセス開始を示す信号を送出する。それとともに、データ線220を通して2つのアドレス、コマンドと、1つのデータを送出する。

10

20

30

40

50



【0103】SMアクセス制御部105は、データ線220を通して送られてきた2つのアドレス、コマンドと、1つのデータをパケットバッファ303に格納する。データ転送制御部310はアービトレーションを行ってパシフ301の使用権を決定し、セクタ302を切り替える。

【0104】図13は、二重化した2つの共有メモリ部15へデータを書き込む場合の、SMアクセス制御部105から2つのSMコントローラ108へのアクセスの流れを示している。2つの共有メモリ部を二重化する場合、一方をマスタ、もう一方をスレーブに設定する。SMアクセス制御部105内のデータ転送制御部310は、アービトレーションによって、まずマスタ側のSMコントローラ108へのアクセスパス2:137の使用権を決定し、制御線6:222によってマスタ側のSMコントローラ108内のデータ転送制御部315へアクセス開始を示す信号(REQ)を出す(ステップ901)。続いて、アドレス、コマンド、及びデータを連続して送出する(ステップ902)。

【0105】マスタ側のSMコントローラ108内のデータ転送制御部315は、制御線6:222によってREQ信号を受け取ると、次にアクセスパス2:137を通して送られてくるアドレス、コマンド、及びデータを受信し、adr、cmd解析部305で解析したアクセス要求に基づいてアービトレーションを行い(ステップ903)、セクタ309を切り替える。データはパケットバッファ303に格納する。アービトレーションの結果、メモリモジュール109へのアクセス権を得たら、メモリの制御情報をメモリ制御部307へ送出し、メモリアクセスのための前処理を行う(ステップ904)。次に、パケットバッファ303からデータを読み出し、セクタ309を介してメモリモジュール109へ書き込む(ステップ905)。

【0106】メモリモジュール109へのアクセスが終了すると、メモリアクセスの後処理を行い、データ転送制御部315においてアクセス状況を示すステータス(STATUS)を生成する(ステップ906)。次に、ステータスをSMアクセス制御部105へ送出する(ステップ907)。

【0107】SMアクセス制御部105はステータスを受け取ったら、マスタ側のSMコントローラ108へのアクセスパス2:137の使用権を開放せずに、アービトレーションによって、スレーブ側のSMコントローラ108へのアクセスパス2:137の使用権を決定する。その後のスレーブ側のSMコントローラ108へのアクセス手順(ステップ908~914)は、マスタ側のSMコントローラ108へのアクセス手順(ステップ901~907)と同様である。

【0108】SMアクセス制御部105内のデータ転送制御部310はスレーブ側のSMコントローラ108か

らステータスを受け取ると、マスタ側とスレーブ側両方のSMコントローラ108へのREQ信号をオフする(ステップ715a、715b)。

【0109】二重化した共有メモリ部15にデータを二重に書き込む場合、上記のように共有メモリ部15をマスタとスレーブに分け、マスタ側、スレーブ側の順でデータを書き込み、スレーブ側へのデータの書き込みが終わるまで両方のアクセスパス2:137を開放しない。これによりデータを書き込む順番が保証され、二重に書き込むデータの一方のデータの書き込み処理を行っている間に、もう一方のデータが他のアクセスパスからの書き込み要求によって書き換えられるのを防ぐことができる。

【0110】SMアクセス制御部105内のデータ転送制御部310はステータスを受け取ると、制御線5:221により、マイクロプロセッサ101へ共有メモリ部15へのアクセスの終了を報告する。

【0111】[実施例5]図14に、実施例1のディスクアレイ制御装置1のチャンネルIF部11、ディスクIF部12、セクタ部13、キャッシュメモリ部14、共有メモリ部15を実装するときの構成を示す。

【0112】チャンネルIF部11、ディスクIF部12、セクタ部13、キャッシュメモリ部14、共有メモリ部15は、それぞれ独立したパッケージ(PK)、すなわち、チャンネルIFPK1、ディスクIFPK2、セクタPK3、メモリPK4にそれぞれ実装する。異なるキャッシュメモリ部14は、異なるパッケージ上に実装する。また、異なる共有メモリ部15も、異なるパッケージ上に実装する。1つのキャッシュメモリ部14と1つの共有メモリ部15は同一のパッケージ上に実装しても問題ない。図14では、1つのキャッシュメモリ部14と1つの共有メモリ部15を同一のメモリPK4上に実装した例を示している。

【0113】チャンネルIFPK1、ディスクIFPK2、セクタPK3、メモリPK4は、ブラッタ5上に実装し、それを筐体に搭載する。

【0114】ここで、各PKをブラッタ5に実装するときには、セクタPK3を中心付近より外側、好ましくはブラッタの両端に配置することが重要となる。このような配置により、各PK間を結ぶ線をブラッタ5上に配線する際、ブラッタ5全体にわたって配線の密度を均一にでき、ブラッタ上の配線を容易になる。

【0115】本実施例では、セクタ部13をセクタPK3に実装して、ブラッタ5の両端に配置するとした。しかし、セクタ部13をパッケージに実装せず、ブラッタの両端に直接実装しても問題ない。

【0116】[実施例6]図1に示す実施例1のディスクアレイ制御装置1において、SMアクセス制御部105とSMコントローラ108間をアクセスパス2:137で接続する代わりに、図7に示すように、2本の共有

バス 130 を介して接続する。SM アクセス制御部 105 からは、2 本の共有バス 130 それぞれに 1 本ずつ接続バスを設ける。また、SM コントローラ 108 からも、2 本の共有バス 130 それぞれに 1 本ずつ接続バスを設ける。SM アクセス制御部 105 及び SM コントローラ 108 は、それぞれ 2 つのアービタを有している。2 つのアービタはそれぞれ 2 本の共有バスのアービトレーション用のアービタである。SM アクセス制御部 105 から SM コントローラ 108 へのアクセスの際は、複数のアービタの内の 1 つがマスタとなり、共有バス 130 の使用権のアービトレーションを行う。そして、使用権を得た SM アクセス制御部 105 が SM コントローラ 108 にアクセスを行う。また、アービタを SM アクセス制御部 105 及び SM コントローラ 108 内に設ける代わりに、独立した回路として共有バス 130 に直接接続しても問題ない。

【0117】共有バス接続では、バスのデータ幅を広げることによりデータ転送速度を上げることが可能であり、共有メモリ部 15 へのアクセス時間を短縮可能である。

【0118】上述したように、図 1 におけるアクセスパス 2:137 のデータ幅は、アクセスパス 0:135 のデータ幅よりも 2 倍以上小さくすることができるので、図 1 のようにチャンネル IF 部 11 及びディスク IF 部 12 と共有メモリ部 15 との間をスター接続（1 対 1 接続）しても、共有メモリ部を実装する LSI のピンネックの問題が生じるケースは少ない。とはいえ、アクセスパス 2:137 の本数が増え過ぎてアクセスパス 2:137 を実装できないという問題が生じる可能性もある。そうした場合、本実施例の共有バス接続が有効となる。

【0119】なお、本実施例では、耐障害性の向上という観点から、セクタ部 13、キャッシュメモリ部 14、及び共有メモリ部 15 をそれぞれ二重化しているが、これらを二重化しなくても、上述の効果をえられることは言うまでもない。

#### 【0120】

【発明の効果】本発明によれば、チャンネル IF 部、ディスク IF 部-キャッシュメモリ間のアクセスパスについてはスループットを高くできる。また、チャンネル IF 部、ディスク IF 部-共有メモリ間のアクセスパスにつ

【図面の簡単な説明】

【図 1】本発明によるディスクアレイ制御装置の構成を示す図。

【図 2】従来のディスクアレイ制御装置の構成を示す図。

【図 3】従来のディスクアレイ制御装置の他の構成を示す図。

【図 4】本発明によるディスクアレイ制御装置内の CM アクセス制御部の構成を示す図。

【図 5】本発明によるディスクアレイ制御装置内の SM アクセス制御部の構成を示す図。

【図 6】本発明によるディスクアレイ制御装置内のセクタ部の構成を示す図。

【図 7】本発明によるディスクアレイ制御装置内のキャッシュメモリ部の構成を示す図。

【図 8】本発明によるディスクアレイ制御装置内の共有メモリ部の構成を示す図。

【図 9】キャッシュメモリ部へのデータの書き込み時の手順を示す図。

【図 10】共有メモリ部へのデータの書き込み時の手順を示す図。

【図 11】二重化した 2 つのキャッシュメモリ部へデータを二重に書き込む時の手順を示す図。

【図 12】1 つのキャッシュメモリ部から別のキャッシュメモリ部へデータをコピーする時の手順を示す図。

【図 13】二重化した 2 つの共有メモリ部へデータを二重に書き込む時の手順を示す図。

【図 14】本発明によるディスクアレイ制御装置内の実装の構成を示す図。

【図 15】本発明によるディスクアレイ制御装置内の他の構成を示す図。

【図 16】本発明によるディスクアレイ制御装置内の他の構成を示す図。

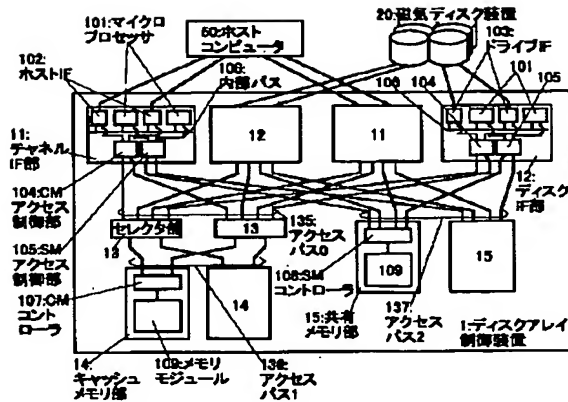
【図 17】本発明によるディスクアレイ制御装置の構成を示す図。

#### 【符号の説明】

1…ディスクアレイ制御装置、11…チャンネル IF 部、12…ディスク IF 部、13…セクタ部、14…キャッシュメモリ部、15…共有メモリ部、20…磁気ディスク装置、50…ホストコンピュータ、101…マイクロプロセッサ、102…ホスト IF、103…ドライブ IF、104…CM アクセス制御部、105…SM アクセス制御部、106…内部バス、107…CM コントローラ、108…SM コントローラ、109…メモリモジュール、135…アクセスパス 0、136…アクセスパス 1、137…アクセスパス 2。

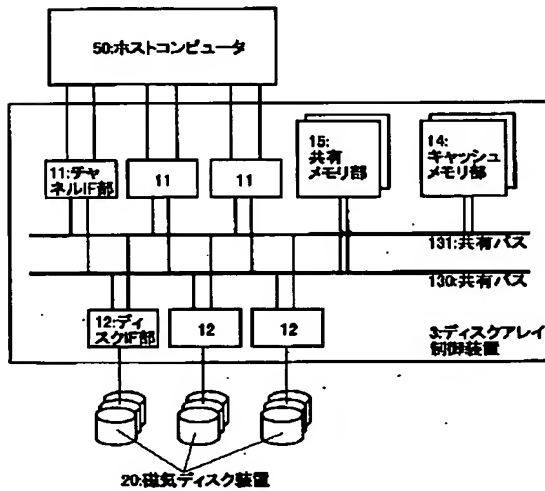
【図1】

図1



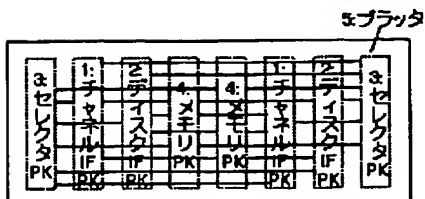
【図3】

図3



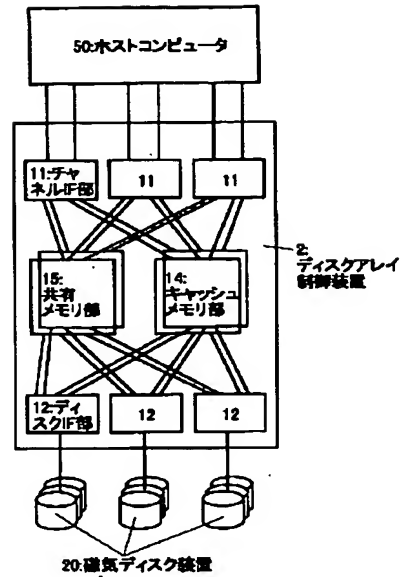
【図14】

図14



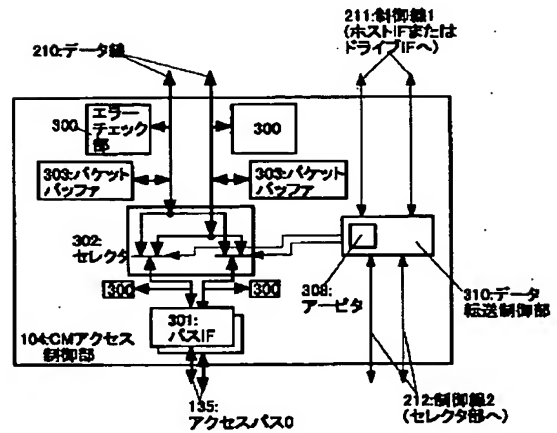
【図2】

図2



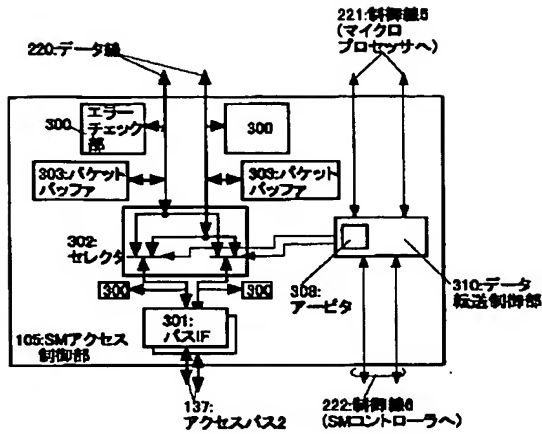
【図4】

図4



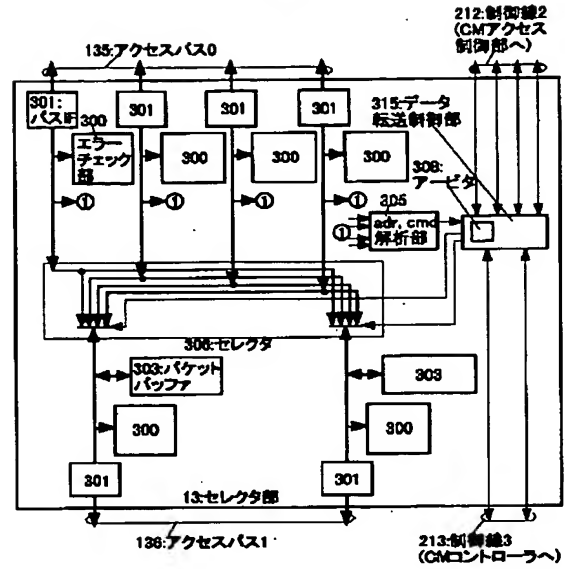
【図5】

図5



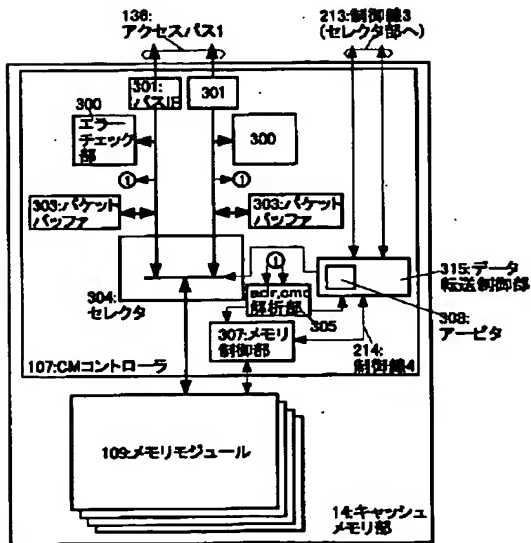
【図6】

図6



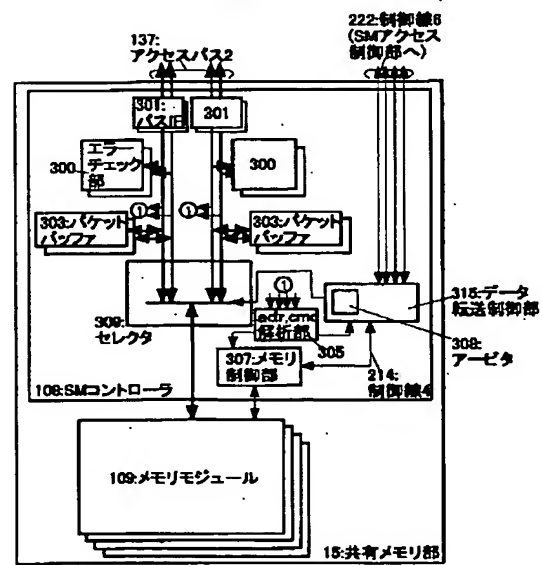
【図7】

図7



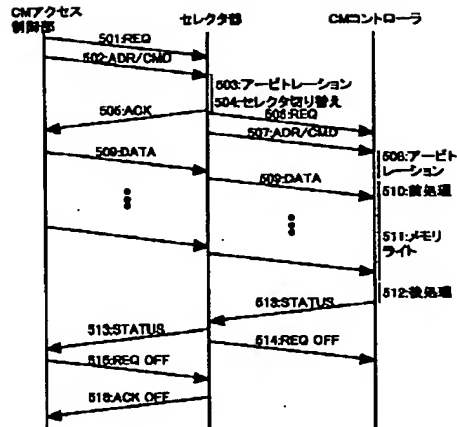
【図8】

図8



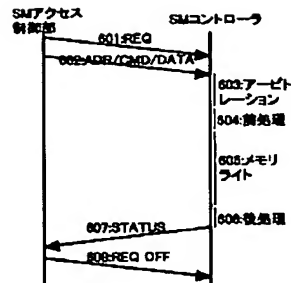
【図9】

図9



【図10】

図10

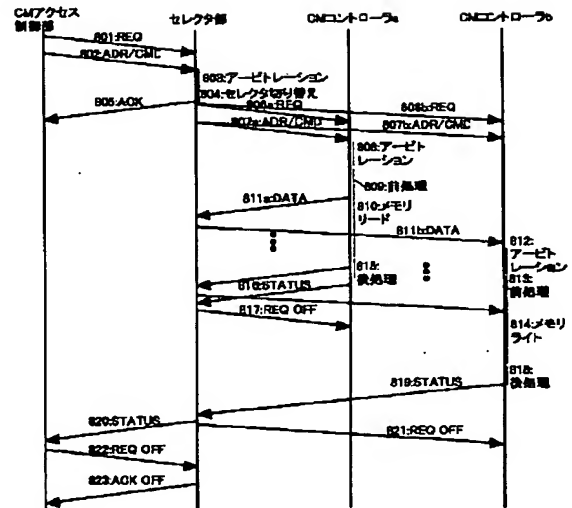
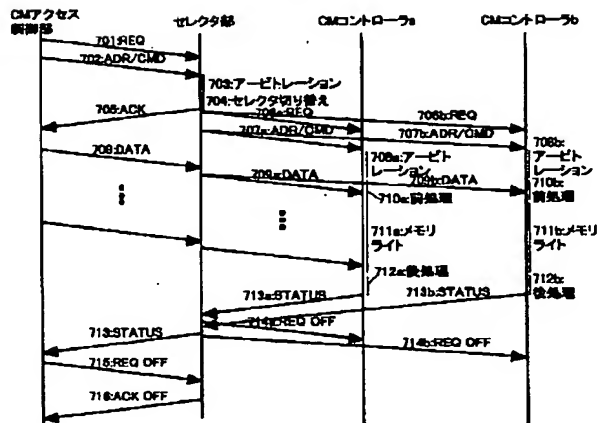


【図12】

図12

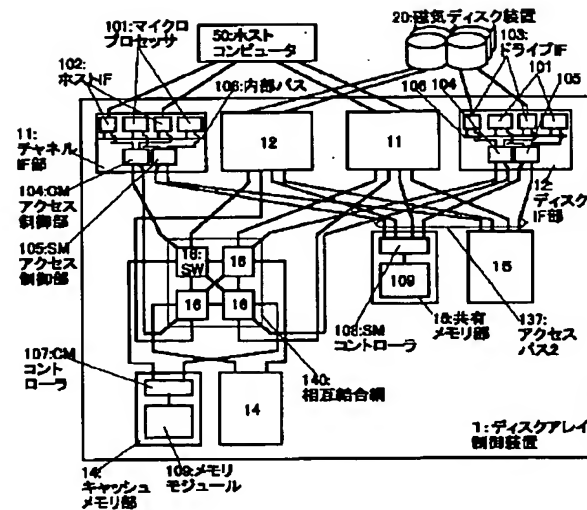
【図11】

図11



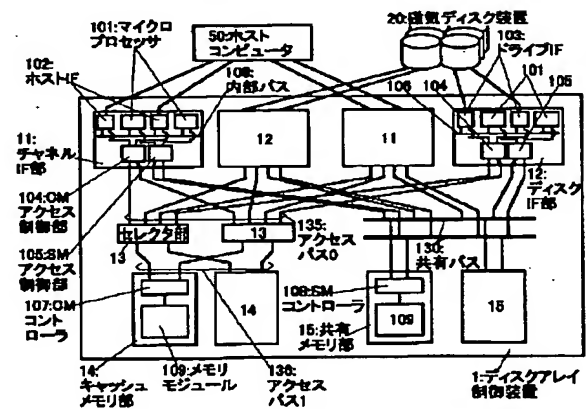
【図 15】

**圖15**



【图 16】

**圖17**



(72)発明者 箕輪 信幸  
神奈川県小田原市国府津2880番地 株式会社  
日立製作所ストレージシステム事業部内  
F ターム(参考) 5B005 JJ01 JJ12 MM12 NN12 WW12  
5B065 BA01 CA11 CA30 CE14 CH01

【公報種別】特許法第17条の2の規定による補正の掲載  
 【部門区分】第6部門第3区分  
 【発行日】平成17年11月17日(2005.11.17)

【公開番号】特開2003-263279(P2003-263279A)  
 【公開日】平成15年9月19日(2003.9.19)  
 【出願番号】特願2003-86908(P2003-86908)  
 【国際特許分類第7版】

G 0 6 F 3/06  
 G 0 6 F 12/08

【F I】

G 0 6 F 3/06 3 0 2 A  
 G 0 6 F 3/06 5 4 0  
 G 0 6 F 12/08 5 1 1 Z  
 G 0 6 F 12/08 5 4 1 Z  
 G 0 6 F 12/08 5 5 7

【手続補正書】  
 【提出日】平成17年9月22日(2005.9.22)  
 【手続補正1】  
 【補正対象書類名】明細書  
 【補正対象項目名】特許請求の範囲  
 【補正方法】変更  
 【補正の内容】  
 【特許請求の範囲】  
 【請求項1】

ホストコンピュータとの複数のチャネルインターフェースと、  
ディスク装置との複数のディスクインターフェースと、  
前記チャネルインターフェースと前記ディスクインターフェースを1対1接続可能な第一  
のアクセスパスと、  
前記チャネルインターフェースと前記ディスクインターフェースを接続する第二のアクセ  
スパスを備え、  
前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータ  
は前記第一のアクセスパスを介して伝送され、  
また、データ伝送に関する制御情報は前記第二のアクセスパスを介して伝送され、さらに  
、前記第一のアクセスパスのビット幅は、前記第二のアクセスパスのビット幅よりも広い  
ことを特徴とするディスクアレイ制御装置。

【請求項2】

ホストコンピュータとの複数のチャネルインターフェースと、  
ディスク装置との複数のディスクインターフェースと、  
前記チャネルインターフェースと前記ディスクインターフェースを1対1接続可能なデー  
タ用ネットワークと、  
前記チャネルインターフェースと前記ディスクインターフェースを接続する制御情報用ネ  
ットワークを備え、  
前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出しされるデー  
タは前記データ用ネットワークを介して伝送され、  
データ伝送に関する制御情報は前記制御情報用ネットワークを介して、前記データ用ネッ  
トワークとは独立に、伝送されることを特徴とするディスクアレイ制御装置。

【請求項3】

ホストコンピュータとの複数のチャネルインターフェースと、



ディスク装置との複数のディスクインターフェースを備え、  
前記チャンネルインターフェースは第一及び第二のアクセス制御部を備え、  
前記ディスクインターフェースは第三及び第四のアクセス制御部を備え、  
前記第一のアクセス制御部と前記第三のアクセス制御部は1対1接続可能であり、  
前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータ  
は、前記第一のアクセス制御部及び前記第三のアクセス制御部を介して伝送され、  
データ伝送に関する制御情報は、前記第二のアクセス制御部及び前記第四のアクセス制御  
部を介して伝送されることを特徴とするディスクアレイ制御装置。

【請求項4】

ホストコンピュータとの複数のチャンネルインターフェースと、  
ディスク装置との複数のディスクインターフェースと、  
前記複数のチャンネルインターフェースと前記複数のディスクインターフェースを接続する  
経路切替手段を備え、  
データ伝送に関する制御情報は、前記経路切替手段を介して伝送されることを特徴とする  
ディスクアレイ制御装置。

【請求項5】

ホストコンピュータとの複数のチャンネルインターフェースと、  
ディスク装置との複数のディスクインターフェースと、  
前記複数のチャンネルインターフェースと前記複数のディスクインターフェースを接続する  
第一の経路切替手段と、  
前記複数のチャンネルインターフェースと前記複数のディスクインターフェースを接続する  
第二の経路切替手段を備え、  
前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータ  
は、前記第一の経路切替手段を介して伝送され、  
データ伝送に関する制御情報は、前記第二の経路切替手段を介して伝送されることを特徴  
とするディスクアレイ制御装置。

【請求項6】

ホストコンピュータとの複数のチャンネルインターフェースと、  
ディスク装置との複数のディスクインターフェースと、  
前記チャンネルインターフェースと第一のアクセスパスを介して接続され、前記ディスクイ  
ンターフェースと第二のアクセスパスを介して接続された経路切替手段を備え、  
前記チャンネルインターフェースは、データ伝送に関する制御情報を前記第一のパスを介し  
て前記複数のディスクインターフェースと通信し、  
前記ディスクインターフェースは、データ伝送に関する制御情報を前記第二のパスを介し  
て前記複数のチャンネルインターフェースと通信することを特徴とするディスクアレイ制御  
装置。

【請求項7】

ディスクアレイ制御装置であって、  
第一の外部装置との第一のインターフェースと、  
第二の外部装置との第二のインターフェースと、  
前記第一のインターフェースと前記第二のインターフェースを1対1接続可能な第一のア  
クセスパスと、  
前記第一のインターフェースと前記第二のインターフェースを接続する第二のアクセスパ  
スを備え、  
前記第二の外部装置に書き込まれるデータ、又は前記第二の外部装置から読み出されるデ  
ータは、前記第一のアクセスパスを介して伝送され、  
データ伝送に関する制御情報は、前記第二のアクセスパスを介して伝送され、  
前記第一のアクセスパスのビット幅は前記第二のアクセスパスのビット幅よりも広いこと  
を特徴とするディスクアレイ制御装置。

【請求項8】

ディスクアレイ制御装置であって、  
第一の外部装置との第一のインターフェースと、  
第二の外部装置との第二のインターフェースと、  
前記第一のインターフェースと前記第二のインターフェースを1対1接続可能なデータ用  
ネットワークと、  
前記第一のインターフェースと前記第二のインターフェースを接続する制御情報用ネット  
ワークを備え、  
前記第二の外部装置に書き込まれるデータ、又は前記第二の外部装置から読み出されるデ  
ータは、前記データ用ネットワークを介して伝送され、  
データ伝送に関する制御情報は前記制御情報用ネットワークを介して、前記データ用ネッ  
トワークとは独立に、伝送されることを特徴とするディスクアレイ制御装置。

【請求項 9】

ディスクアレイ制御装置であって、  
第一の外部装置との第一のインターフェースと、  
第二の外部装置との第二のインターフェースと、  
前記第一のインターフェースと前記第二のインターフェースを接続する第一の経路切替手  
段と、  
前記第一のインターフェースと前記第二のインターフェースを接続する第二の経路切替手  
段と、  
前記第二の外部装置に書き込まれるデータ、又は前記第二の外部装置から読み出されるデ  
ータは前記第一の経路切替手段を介して伝送され、  
データ伝送に関する制御情報は前記第二の経路切替手段を介して伝送されることを特徴と  
するディスクアレイ制御装置。

【請求項 10】

ディスクアレイ制御装置であって、  
第一の外部装置との複数の第一のインターフェースと、  
第二の外部装置との複数の第二のインターフェースと、  
前記複数の第一のインターフェースと第一のアクセスパスを介して接続され、前記複数の  
第二のインターフェースと第二のアクセスパスを介して接続された経路切替手段を備え、  
前記第一のインターフェースは、データ伝送に関する制御情報を前記第一のアクセスパス  
を介して、前記複数の第二のインターフェースと通信し、  
前記第二のインターフェースは、データ伝送に関する制御情報を前記第二のアクセスパス  
を介して、前記複数の第一のインターフェースと通信することを特徴とするディスクアレ  
イ制御装置。

【請求項 11】

ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部  
と、  
ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、  
前記チャネルインターフェース部と前記ディスクインターフェース部を1対1接続可能な  
第一のアクセスパスと、  
前記チャネルインターフェース部と前記ディスクインターフェース部を接続する第二のア  
クセスパスを備え、  
前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータ  
は、前記第一のアクセスパスを介して伝送され、  
データ伝送に関する制御情報は、前記第二のアクセスパスを介して伝送され、  
前記第一のアクセスパスのビット幅は前記第二のアクセスパスのビット幅よりも広いこ  
とを特徴とするディスクアレイ制御装置。

【請求項 12】

ホストコンピュータとのインターフェースを有する複数のチャネルインターフェース部  
と、

ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、  
前記チャンネルインターフェース部と前記ディスクインターフェース部を1対1接続可能な  
データ用ネットワークと、  
前記チャンネルインターフェース部と前記ディスクインターフェース部を接続する制御情報  
用ネットワークを備え、  
前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータ  
は、前記データ用ネットワークを介して伝送され、  
データ伝送に関する制御情報は前記制御情報用ネットワークを介して、前記データ用ネッ  
トワークとは独立に、伝送されることを特徴とするディスクアレイ制御装置。

【請求項13】

ホストコンピュータとのインターフェースを有する複数のチャンネルインターフェース部  
と、  
ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、  
前記チャンネルインターフェース部と前記ディスクインターフェース部を接続する第一の経  
路切替手段と、  
前記チャンネルインターフェース部と前記ディスクインターフェース部を接続する第二の経  
路切替手段と、  
前記ディスク装置に書き込まれるデータ、又は前記ディスク装置から読み出されるデータ  
は前記第一の経路切替手段を介して伝送され、  
データ伝送に関する制御情報は前記第二の経路切替手段を介して伝送されることを特徴と  
するディスクアレイ制御装置。

【請求項14】

ホストコンピュータとのインターフェースを有する複数のチャンネルインターフェース部  
と、  
ディスク装置とのインターフェースを有する複数のディスクインターフェース部と、  
前記複数のチャンネルインターフェース部と第一のアクセスパスを介して接続され、前記複  
数のディスクインターフェース部と第二のアクセスパスを介して接続された経路切替手段  
を備え、  
前記チャンネルインターフェース部は、データ伝送に関する制御情報を前記第一のアクセ  
スパスを介して、前記複数のディスクインターフェース部と通信し、  
前記ディスクインターフェース部は、データ伝送に関する制御情報を前記第二のアクセ  
スパスを介して、前記複数のチャンネルインターフェース部と通信することを特徴とするディ  
スクアレイ制御装置。